Mønness, E. (2012). Data Transformations with a Full 2^6 Experimental Design—A Metal-Cutting Case Study. *Quality Engineering 24*(1), 37-48.
DOI: http://dx.doi.org/10.1080/08982112.2011.616150

DATA TRANSFORMATIONS WITH A FULL 2^6 EXPERIMENTAL DESIGN A METAL CUTTING CASE STUDY

E. Mønness

Hedmark University College, Rena, Norway

THIS IS A PREPRINT OF

E. Mønness (2012): Data Transformations with a Full 2^6 Experimental Design—A Metal-Cutting Case Study,

Quality Engineering, 24:1, 37-48

To link to this article: http://dx.doi.org/10.1080/08982112.2011.616150

Abstract
The Box-Cox transformation was evaluated with reference to a six-factor full factorial (2^6) data set with 64 runs. The data were used to determine the optimal operating conditions for a milling machine with respect to surface finish. A suitable transformation was determined by minimizing the mean square errors, evaluating the size of the effect significances, the normal probability plots of the estimated effects, Shapiro-Wilk test and the model residuals. The achievement of both normality with constant variance and a simple model came about as a result of a trade-off between several different criteria.

Key words: Design of experiments, Box-Cox transformations, Normal probability plot, Residuals, Expectations.

1. Introduction

The traditional reasons for making a nonlinear transformation are (Box and Cox 1964):

- R1. To make a model simpler. This often means finding a scale such that the model parameterization has fewer interactions.
- R2. To achieve homoscedasticity, i.e. equal variances throughout the experiment.
- R3. To ensure that the data have normal distribution of errors.

Box and Cox (1964) introduced the following set of transformations for Y>0:

$$Y^{\{\lambda\}} = \begin{cases} \dfrac{Y^{\lambda}-1}{\lambda}\,\dot{Y}^{(1-\lambda)} & \text{for} \quad \lambda \neq 0 \\ \log(Y)\,\dot{Y} & \text{for} \quad \lambda = 0 \end{cases}$$

where $\dot{Y}$ is the geometric mean of the data concerned. The factor $\dot{Y}^{(1-\lambda)}$ is optional but ensures that the residual sums-of-squares from different λ are comparable (i.e. on the same scale). The use of this well-established procedure is described herein, even though it has been called into question (Dagenais and Dufour 1994).

The aim of this paper is to present a case study evaluating the Box-Cox transformation on a data set with several active factors and a large set of possible effects. It makes it possible to analyse the data with normal plots and residual analysis and demonstrate the trade-offs involved in deciding a proper model for the data. The most common violation of R2, at least in industrial experiments, is that variation increases with the expectation. Making plots of residuals against predicted values is an easy and good method of assessing this issue. Normal plots are a good method of assessing normality (R3), but the Shapiro-Wilk (SW) test (Shapiro and Wilk, 1965) is also available.

In the Biological 3x4 experiment with replicates, (Box, Hunter et al. 2005) R1 and R2 appear to have been reached simultaneously using the same transformation. In the textile 3^3 experiment without replicates (Box and Cox 1964), and the 2^4 without replicates drill example (Daniel 1976), revisited in Box, Hunter et al. (2005), R1 is the criterion of interest, and a log-likelihood method is used, assuming R2 and R3.
However, a transformation achieving R1 might be different from a transformation achieving R2 and R3 (Nelder 1968). A sequential search might be performed (Box and Cox 1964). R2 is more important than R3, because the least squares estimates are unbiased and independent of the error distribution. Hinkelmann and Kempthorne (1994) argue that in a completely

randomized design, with additivity between treatments and experimental variation, analysis of variance (ANOVA) with replicates can be tested distribution-free by means of a permutation test, and that the common F test is a good approximation for this. Thus, the use of a transformation is irrelevant. However, additivity between treatments and experimental variation are crucial, and unequal variances can be an indication of non-additivity in this sense. Under this framework, the use of a transformation could be appropriate for finding a scale with R2 that supports additivity.

When $\lambda<1$ large data values are decreased and right skewed data can be normalized (R3). When $\lambda>1$, left skewed data can be normalized. The practical range of λ values can be said to be from about -3 to +3. Sleeper (2005) advocates λ between -5 and +5. When $|\lambda|$ becomes large, the transformation tends toward a vertical and a horizontal line that have a junction at (1,0) in the $\left(Y \times Y^{\{\lambda\}}\right)$ plane leaving the data with very little variation. $Y^{\{\lambda\}}$ can only be normal in the log() case (and when $\lambda=1$), else it will be a truncated normal (Freeman and Modarres 2006). When $\lambda>0$ then $Y^{\{\lambda\}} > -\lambda^{-1}$ and when $\lambda<0$ then $Y^{\{\lambda\}} < -\lambda^{-1}$. Mønness (2011) presents a use of the inverse Box-Cox transformation that also explores what kind of distributions can be approximately normalized by the Box-Cox transformation.

Model parameters and λ should be estimated simultaneously by maximum likelihood. The likelihood is based on the distribution density

$$f(y) = \frac{y^{\lambda-1}}{\Phi\left(\operatorname{sign}(\lambda)k\right)\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{y^{\lambda}-(\lambda\mu+1)}{\sigma\lambda}\right)^2} \quad , \; y \geq 0$$

Where $k = \frac{\lambda\mu+1}{\lambda\sigma}$ and $\Phi(\;)$ is the standard cumulative normal distribution, (μ,σ) is the expectation (with the model structure) and the standard deviation of the transformed variable. This distribution is called the power-normal distribution (Freeman and Modarres 2006). The shape of the density function is given by λ and k (Goto and Inoue 1980). If $\lambda\leq 0$ or $\lambda\geq 1$ the density is unimodal. A distribution with $k \leq 2\sqrt{\lambda^{-1}-1}$ & $\lambda>0$ has a density with a maximum at x=0. If $k > 2\sqrt{\lambda^{-1}-1}$ & $\lambda<1$ the density has a local minimum close to x=0, but the size of the local minimum is negligible with most k found in practice and the shape appears to be unimodal. Mønness (2011) found that in most practical cases, $\lambda<1$ yields right skewed distributions while $\lambda>1$ yields left skewed distributions although this does not hold generally. Formulas for moments are complicated. Formulas for special cases are given by Freeman and Modarres (2006). If $\lambda<0$, moments of higher order than $|\lambda|$ do not exist (Goto and Inoue, 1980). If the original data can be described by this distribution, the transformed data can be described as normal.

The truncation factor $\Phi\left(\operatorname{sign}(\lambda)k\right)$ is often ignored in the estimation. The estimation is often simplified further by estimating the model expectation with several λs and finding the λ that maximizes the likelihood. Minimizing the mean square errors is equivalent to maximizing the log-likelihoods when the truncation factor is ignored (Box and Cox 1964). Figure 3 below can thus be viewed as an approximate maximum likelihood estimation.

With a specific set of data, there is no guarantee that the minimizing λ also ensures the error distribution to be normal (R3) and with equal variances (R2), this has to be evaluated after the estimation has been done. The minimizing λ can be model dependent, as will be seen.

Transformation is a practical tool to achieve a reasonable sound analysis in real situations and involve several tradeoffs were graphical displays play a critical role.

In econometrics, the untransformed data is sometimes modeled as having a two-parameter gamma distribution (Amemiya, 1981)

The analysis described below was carried out using SYSTAT (Systat Software 2004).

2. The Experiment

The data were originally collected as part of an investigation to compare the techniques of full factorial and the Taguchi methodology (Garzon 2000). The main objective of the study was to optimise a milling process (a metal cutting operation) by determining the significant factors and interactions that affected the surface finish of the metal. Mills and Redford (1983) highlight the factors that influence the result of a milling process. Maekawa (1998) discusses the intricate nature of the relationships involved between the factors that influence the smoothness of the final surface. Prior to the initial investigation using the full factorial experiment, a brainstorming session was held in order to identify the principal factors involved in the milling operation. "*Several factors must be considered when setting up a milling job, including (among many) the type of milling operation, speeds feeds, depth of cut, and safety. There is a very dense network of relationships among factors, which in most cases interact having high incidence on the process itself*" (Garzon 2000). The full factorial experiment involved six factors, each of which had two possible values; hence, a total of 64 runs were required. Full randomization was reported to have occurred. Table 1 shows the selected factors (A, B, C, D, E and F), their levels, and their coded values. The surface roughness was measured using the "centre line average"(Kalpakjian 1992). The values are shown in Table 2, in which a small value represents a smooth surface. For a further discussion of the means by which the predictors and levels were selected, see Garzon (2000).

Mønness, Linsley et al. (2007) made use of the same set of data in an earlier paper on the different fractions of a factorial design. The data used here resembles the drill example of 2^4 without replicates (Daniel 1976).

Industrial experiments are often performed using a fractional factorial design, due to the large numbers of runs involved in a full factorial. The data afford us with the opportunity to test the use of the Box-Cox transform, in combination with a large set of estimable parameters. It would have been preferable also to have had replicates, but the availability of such data would have been even less likely. However, eight repetitions (i.e. measurements using the same piece of treated material) were performed of each treatment combination (run). Of the total (8 x 64), nine measures were missing, all from runs having large surface roughness values. While not actually replications, the repetitions provide some insight into the variability inherent in the data.

Let $Y_{\text{run j}} = \mu_{\text{run}} + \varepsilon_{\text{run}} + \delta_{\text{run j}}$ where run=1,…,64 and *j* is the replicate index.

μ_{run} is the expectation, ε_{run} is the inter-run variation with standard deviation $\sigma(\varepsilon_{\text{run}})$ and $\delta_{\text{run j}}$ is the intra-run variation with standard deviation $\sigma(\delta_{\text{run}})$.

There is a well-established procedure for disentangling location and dispersion using a power transformation of the data. Let δ be the replicate variation (within the experimental unit). If $\sigma(\delta_{\text{run}}) \propto \text{mean}_{\text{run}}^{(1-\lambda)}$ then Y^{λ} is a transformation that is suitable for ensuring homoscedasticity

(Box, Hunter et al. 1978). Figure 1a shows the relationship between $\mathrm{Log}_{10}\left(\sigma\left(\delta_{run}\right)\right)$ and $\mathrm{Log}_{10}\left(\overline{Y_{run\bullet}}\right)$ for all 64 runs. The linear regression having a slope of 1.78 is superimposed on the figure. The transformation using the reciprocal Y^{-1} yields the data shown in Figure 1b. This is not quite the same as the $Y^{-0.8}$ suggested by the regression, but it is a natural approximation. The linear regression having a slope of 0.18 is superimposed on the figure. Because $\sigma\left(\delta_{run}\right)$ is the intra-unit variation, it only affords us an indication of the nature of the inter-unit variation. The repetitions are not replicates (Wu and Hamada 2000)p9. Some theory of transformations for mixed models exists (Gurka, Edwards et al. 2005), however, this is not made an issue herein. We herein restrict our attention to

$\overline{Y_{run\bullet}} = \mu_{run} + \varepsilon_{run}$, i.e. the means for each run.

ε_{run} is the random error between runs, which now includes all the sources of variation, including design variation and the means of the intra-unit variation. $\sigma\left(\varepsilon_{run}\right)$ should be stable across runs (R2), and ideally ε_{run} should be normally distributed (R3).

The data are shown in Figure 2a, and given in detail in Table 2. Most runs yield a smooth surface with small values while some runs produce results that are less good. The residuals from the model $\mu_{run} = \mu + A + B + C + D + E + F$ + all the two-factor and three-factor interactions are shown in Figure 2b. The regression is dominated by a few large observations and the residuals appear neither to have equal variances nor to be normally distributed. The data are right-skewed. In the data, the value of the maximum divided by the minimum is 25.56/0.78 = 32.8. A transformation seems appropriate and could make a difference.

3. Results

Since the original data is right skewed, λs less than one might normalize the transformed data. Five sets of regressions are made for λ = -3 to +2 in steps of 0.1. Values greater than one are included so that the original data will not be a border case. Figure 3 shows the mean square errors (MSE) for different values of λ. (The λ range not shown yields higher MSEs.) The five regression models are:

1. Main effects only, shown as “MSE (main effects only)”, which have the highest MSE values with a minimum that occurs at λ =-1.5.

2. Main effects with the two-factor interactions. The minimum MSE value occurs at λ = -1.2

3. Main effects with both the two- and three-factor interactions The minimum MSE value occurs at λ = -0.3

4. A stepwise regression that allows for the main effects with the two- and three-factor interactions, shown as “MSE (stepwise)”. The stepwise procedure allows for factors to be both entered and removed (Systat Software 2004). This curve has varying degrees of freedom as λ vary, which makes it less smooth. The minimum MSE value occurs at λ = -0.3

5. $\mu_{run} = \mu + A + B + C + D + E + F + DE + DF + EF + DEF$, shown as "MSE (A B C D E F DE DF EF DEF)" The minimum MSE value in this case occurs at λ = -0.85.

Regressions 1, 2, and 3 include all the factors with an increasing degree of complexity. The effect of the stepwise regression 4 is to simplify regression 3. Regression 5 was considered because the interactions between D, E, and F turned out to be the most influential.
It may be seen that when the complexity of the model is increased, the mean square error is decreased (as is often the case), but in our case the proposed value of λ increases from -1.5 to -0.3, thus λ is model dependent . The quotient obtained between the curves (main, two-factor and three-factor interactions for each value of λ) represent a series of F tests in which some interactions are significant.

Table 3 shows the full ANOVA output for regression 3 on the untransformed (λ = 1) data and the reciprocal transformed data (λ=-1). All these values can also be seen in Figure 4 and 5. Figure 4 shows t-values obtained from regression 3 for the λ values . The scale on the vertical axis should be noted. The horizontal dashed lines are the 5 % significance limits of a single t test.

It may be seen that:

- D and E are the most important factors, regardless of the choice of λ. The next most important factor is F. A and B are never significant.
- The interactions between D, E, and F are highly significant.
- C, and some other effects, seems significant when λ is close to zero.
- The value of λ that gives the highest significances for the seemingly most important factors also gives high significances for several other factors, indicating a complex model violating R1.

Testing 6 main effects, 15 two- and 20 three-factor interactions with regression model no. 3 is a multiple test situation. Some of the significances found in Figure 4 may be due to multiple testing issues. The estimated effects are orthogonal, and therefore uncorrelated. They also have equal estimates of variance. They may thus be judged using a normal probability plot. In Figure 5, in which each of the four panels depicts a normal probability plot for a different value of λ, the estimates are divided by their common standard deviation; thereby making the values on the horizontal axis equal to t-values (The division changes the scale, not the pattern).
Figure 5 shows that:

- For λ = 1, the effects hardly constitute a set of normal variates with expectations equal to zero, with a few non-zero expectations (thereby providing further support for a transformation).
- Figure 4 depicts a large number of significances when λ = 0. However, the probability plot in Figure 5b indicates that this may be due to a normal variation among several values that have an expectation of zero (i.e. not significant).
- The significance of C when λ = 0 is dubious. Its significant value is close to being just a maximum among several values with an expectation of zero. At λ = -1, C is within the set of normal variates judged to have an expectation of zero. This is also the case when λ = -0.5 (not shown).
- The extreme value λ = -3 presents a picture where only D and DE (maybe E) are significant.

The residuals that pertain after the fit has been applied should be normally distributed (R2), and should have a random pattern with equal variation (R3) against the fitted value. This is shown for λ = 0, -1 and -3 in Figure 6. The frames e and f for λ=-3 seems to best fit normal distribution properties: no extreme residuals and very little variation structure against the fitted value. However, the transformation may have taken to much variation out of the data and λ=-3 is certainly not advised by the minimizing MSE as shown in Figure 3. The discrimination between λ = 0 and λ = -1 is not very strong, but may be judged in favour of λ = -1, in that the residuals appear slightly more normal (R3) and the variation depends slightly less on the predicted values (R2).
Normality can be examined formally by the Shapiro-Wilk (SW) test. The more significant, the less normal is the data. The significances of Shapiro-Wilk tests for every λ are shown in Figure 7for both regression model 3 and 5. With model 3 SW points to λ=0 as the best option, while with model 5 SW points to λ=-1.

We judge that the combination (D=+1, E=-1, F=+1) represents the best treatment, alternatively (D=+1, E=-1, F=-1) (See Table 4). A, B, and C can be set according to external experimental preferences. Maybe C should be scrutinised in a follow-up experiment.

4. The Expected Values

Table 4 shows the estimated expected values, on their original scale, for experimental settings in which E, D, and F are varied (A, B, and C are artificially set to zero), for λ = -1, 0 and +1. For λ = -3 also F is removed. The error degrees of freedom for the confidence interval are 64-(1+6+15+20) = 22, even if only D, E, and F and their interactions (eight model degrees of freedom, four when λ = -3) are in action. (The number of estimated factors will only affect the error and the error degrees of freedom because the effects are orthogonal. I.e. the interval is conservative).
The composite estimates are nearly equal, regardless of the value of λ, except for λ=-3. The precision of the estimates at small values are improved by a transformation. Due to the non-linearity, the confidence intervals on the original scale are asymmetric and vary in size.
Some basic statistical theory is required to highlight the issue of expectation. Consider the general regression model in matrix notation (Searle 1971)

$\mathbf{y} = \mathbf{Xb} + \mathbf{e}$ with $\mathrm{E}(\mathbf{y}) = \mathbf{Xb}$.

$\mathbf{y}$ is a n*1 vector of observations, $\mathbf{X}$ is a n*p known design matrix of full row rank, $\mathbf{b}$ is a p*1 vector of the effects to be estimated and $\mathbf{e}$ is a n*1 vector of uncorrelated random errors.
The estimated expectations at the design points are

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\mathbf{b}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

If the fit is good, $\hat{\mathbf{y}} \approx \mathbf{y}$. Now, consider $\mathbf{y}$ as a nonlinear monotone transformation $\mathbf{y} = f(\mathbf{Y})$ where $\mathbf{Y}$ represents the original observations, $f(\mathbf{Y})$ is the Box-Cox transformation for some λ, and $f^{-1}(\ )$ is the inverse transformation. Then the estimated expected values on the original scale are

$$\mathrm{f}^{-1}\left(\widehat{\mathrm{f}(\mathbf{Y})}\right) \approx \mathrm{f}^{-1}(\mathrm{f}(\mathbf{Y})) = \mathbf{Y}\ .$$

λ = -3 is an exception for higher estimated values. As mentioned in the introduction the transformation becomes very flat for high $|\lambda|$ values. Thus the inverse transformation becomes inaccurate and might even be impossible due to the truncation issue.

In industrial experiments, the fit is often good, both because the value of the "error" term is actually small, and because the number of runs (observations) is minimized. If p is close to n, the fit is typically good due to pure linear algebra. In any case, a reasonable nonlinear transformation will not alter the estimated expected values that much. Its aim is to assist in the selection of columns that constitute the final matrix $\mathbf{X}$.

5. Discussion and conclusion.

The search for an appropriate transformation involves meeting criteria R1, R2, and R3 simultaneously. In the case study described herein, R2 and R3 were reasonable satisfied using the Box-Cox transformation. These two criteria were obtained together, as seen in Figure 5 and Figure 6. In the case of R1, however, the author sees no special reason that simplicity of expectation (R1) should appear together with R2 and R3, or the opposite. R2 and R3 must be obtained *before* standard normal theory tests can be applied in search of a simple model (R1). The selection of λ by minimizing MSE in this case was dependent on the model in question. The most complex model obtained pointed to $\lambda \approx 0$. R2 and R3 are reasonably clearly satisfied for $\lambda = 0$ (Figure 6). However, this transformation gives rise to a large set of significant factors (see Figure 4 and 5). The use of $\lambda = -1$ still yielded R2 and R3, however simplified the model. Effect C was significant on the original scale, but not when $\lambda = -1$. This shows that achieving R1, R2 and R3 together may well be a considerable challenge when the model itself is complicated. Using a formal test for normality does not rule out the need of assessing the practical situation. With regression model 3, the Shapiro-Wilk test points to $\lambda = 0$ as giving the the most normal data. (Figure 7), while model 5 is rejected. When $\lambda = 0$ there is a large set of significant effects that will be embraced in the residuals when using model 5, explaining the SW rejection of normality in this case. With regression model 5, the SW test points to $\lambda = -1$. Note that both Figure 6 and the SW points to $\lambda = -3$ as good candidate (while the minimizing of MSE does not!). The residuals from $\lambda = -3$ might be normal, but too much structure are removed from the data in this case.
A combination of transformation, normal probability plots, formal tests and residual plots can assist in the trade-off decision for the use of a particular model.

The transformation is crucial to the establishment of R2 and R3 used in selecting the design matrix $\mathbf{X}$. The criterion R1 involves finding the smallest set of $\mathbf{X}$ columns that can "recreate" the data with some restrictions such as the heredity principle (Wu and Hamada 2000, p112): "*In order for an interaction to be significant, at least one of its parent factors should be significant*". Furthermore, the precision of the estimates obtained with small values was improved by carrying out the transformation. This is beneficial in itself. When an $\mathbf{X}$ is decided, the estimated response surface, if with reasonable fit, will be found to be stable independent of the transformation.

References

Amemiya, T., Powell, J. L. (1981) A comparison of the Box-Cox maximum likelihood estimator and the non-linear two-stage least squares estimator. Journal of Econometrics, Volume 17, Issue 3, December 1981, Pages 351-381

Box, G., J. Hunter, et al. (2005). Statistics for Experimenters Design, Innovation and Discovery John Wiley & Sons.

Box, G. E. P. and D. R. Cox (1964). An Analysis of Transformations. Journal of the Royal Statistical Society, Series B **26**(2): 211-252.

Box, G. E. P., W. G. Hunter, et al. (1978). Statistics for Experimenters, John Wiley & Sons.

Dagenais, M. G. and J. M. Dufour (1994). PITFALLS OF RESCALING REGRESSION-MODELS WITH BOX-COX-TRANSFORMATIONS. Review of Economics and Statistics **76**(3): 571-575.

Daniel, C. (1976). Applications of statistics to industrial experimentation. New York, Wiley.

Freeman, J., and Modarres, R. (2006). Inverse Box-Cox: The power-normal distribution. Stat. Probab. Lett. 76(8): 764-772.

Garzon, I. (2000). Optimisation for Product and Process Improvement: Investigation of Taguchi Tools and Genetic Algorithms. ISRU. Newcastle Upon Tyne, University of Newcastle Upon Tyne. **Ph.D.**

Goto, M., and Inoue, T. 1980. Some properties of the Power Normal Distribution.. Bulletin Biometric Society of Japan 1: 28-54.

Gurka, M. J., L. J. Edwards, et al. (2005). Extending the Box-Cox transformation to the linear mixed model. Journal of the Royal Statistical Society, Series A **169**(2): 15.

Hinkelmann, K. and O. Kempthorne (1994). Design and Analysis of Experiments. New York, John Wiley & Son.

Kalpakjian, S., Ed. (1992). Manufacturing, Engineering & Technology 2.nd ed. Reading, Mass, Addison Wesley.

Maekawa, K. (1998). Computational aspects of tribology in metal machining Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology **212**(4): 307-318. (10.1243/1350650981542128)

Mills, B. and A. H. Redford (1983). Machinability of engineering materials / B. Mills and A.H. Redford. London :, Applied Science.

Mønness, E., M. Linsley, et al. (2007). Comparing different fractions of a factorial design: A metal cutting case study. Applied Stochastic Models in Business and Industry **23**: 117-128.

Mønness, E. (2011). The Power-Normal Distribution: Application to forest stands. Canadian Journal of Forest Research. 41(4) 707-714.

Nelder, J. A. (1968). Weighted Regression, Quantal Response Data, and Inverse Polynomials. Biometrics **24**(4): 979-985.

Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality. Biometrica, 52, 591-611.

Searle, S. R. (1971). Linear Models. New York, John Wiley & Sons.

Sleeper, A. (2005) Design for Six Sigma Statistics. New York, McGraw-Hill Professional.SYSTAT (2004) SYSTAT 11.0 Chicago, IL USA http://www.systat.com

Wu, J. and M. Hamada (2000). Experiments Planning, Analysis, and Parameter Design Optimization. New York, Wiley.

			Coded	
Predictor	**Low**	**High**	**Low**	**High**
A - Tool speed (rev/min)	2700	3200	-1	+1
B - Workpiece speed (mm/min)	203	330	-1	+1
C - Depth of cut (mm)	0.5	1.0	-1	+1
D - Coolant	Off	On	-1	+1
E - Direction of cut	Conventional	Climbing	-1	+1
F - Number of cuts	1	2	-1	+1

Table 1. The 6 experimental factors

A	B	C	D	E	F	r1	r2	r3	r4	r5	r6	r7	r8
1	-1	1	-1	1	-1	1.080	1.063	0.907	0.930	1.017	1.020	0.947	0.940
1	1	1	1	1	1	0.913	0.883	0.890	0.890	1.023	1.023	1.003	0.987
-1	1	1	-1	-1	1		32.500			15.200	16.200	36.700	27.200
1	-1	-1	-1	-1	1	22.000		15.200	11.700	15.000	12.500	16.500	8.200
1	1	-1	1	-1	-1	1.117	0.847	0.897	0.797	0.767	0.773	0.783	0.840
-1	1	-1	-1	1	-1	1.090	1.020	0.970	1.050	1.050	1.070	1.020	1.050
-1	-1	-1	1	1	1	1.240	1.140	1.090	1.150	1.120	1.110	1.060	1.110
-1	-1	1	1	-1	-1	1.090	1.050	0.950	0.940	1.100	1.120	1.060	1.120
-1	1	-1	1	-1	1	0.930	0.930	0.970	0.940	0.980	0.967	1.013	0.923
-1	-1	-1	-1	-1	-1	1.207	1.547	1.410	2.013	1.490	2.150	1.367	1.637
1	-1	-1	1	1	-1	1.113	1.063	0.973	0.983	1.057	0.987	0.973	0.963
1	-1	1	1	-1	1	0.803	0.780	0.757	0.750	0.767	0.743	0.830	0.803
1	1	-1	-1	1	1	1.217	1.130	1.147	1.155	1.300	1.183	1.247	1.130
1	1	1	-1	-1	-1	1.690	1.597	1.680	2.230	4.407	2.657	3.600	5.033
-1	-1	1	-1	1	1	0.890	0.820	0.910	0.970	0.920	0.860	1.160	1.050
-1	1	1	1	1	-1	0.890	0.850	0.800	0.880	0.900	0.890	0.860	0.880
1	-1	1	-1	1	1	1.223	1.190	1.180	1.177	1.140	1.173	1.180	1.123
-1	1	-1	-1	1	1	1.203	1.118	1.190	1.257	1.173	1.123	1.120	1.140
-1	-1	1	1	-1	1	0.990	0.947	0.967	0.980	0.967	0.953	0.947	1.017
1	-1	-1	-1	-1	-1	1.470	1.407	1.330	1.147	2.690	2.363	2.323	2.147
-1	-1	-1	1	1	-1	1.063	0.990	0.973	0.940	1.090	0.957	0.940	0.917
1	1	1	1	1	-1	0.940	1.040	1.023	1.030	0.897	0.873	0.913	0.980
1	1	-1	1	-1	1	0.990	1.110	1.100	1.130	0.950	0.940	1.010	0.960
-1	1	1	-1	-1	-1	3.080	3.530	4.030	4.400	3.970	4.000	3.670	5.830
-1	-1	1	-1	1	-1	1.037	0.963	1.003	0.990	0.930	0.973	1.013	0.940
-1	-1	-1	-1	-1	1	7.533	6.400	7.900	5.300	1.670	6.233	7.000	11.167
-1	1	1	1	1	1	1.100	1.110	0.963	1.040	1.023	1.017	0.983	1.003
1	1	1	-1	-1	1		30.700		33.000		15.500	15.200	
1	1	-1	-1	1	-1	1.020	1.007	1.007	0.953	0.997	0.990	1.013	1.003
1	-1	1	1	-1	-1	1.040	1.157	1.097	1.030	0.913	0.940	0.940	0.987
1	-1	-1	1	1	1	0.973	0.957	0.930	0.950	0.957	0.923	1.023	1.107
-1	1	-1	1	-1	-1	1.013	1.090	1.117	1.140	1.130	1.217	1.130	1.070
1	1	1	1	-1	1	0.947	0.923	0.967	0.887	0.870	0.907	0.930	0.897
-1	-1	-1	1	-1	1	0.957	0.913	0.897	0.930	0.897	0.887	0.903	0.920
-1	-1	1	1	1	-1	0.953	0.957	1.030	1.020	0.980	0.973	0.983	0.933
-1	1	1	-1	1	1	1.050	1.033	1.073	1.010	0.980	1.073	1.030	1.060
1	-1	1	-1	-1	-1	2.507	2.230	2.080	3.933	2.080	3.013	5.100	5.100
-1	1	-1	-1	-1	-1	1.447	1.340	1.523	1.290	1.440	1.503	1.423	1.420
1	1	-1	1	1	-1	0.960	1.000	1.010	1.000	1.010	0.940	0.910	0.950
1	-1	-1	-1	1	1	1.390	1.570	1.600	1.120	1.030	1.240	1.550	1.240
-1	-1	-1	-1	1	-1	1.053	1.030	1.057	1.063	1.030	1.037	1.013	1.050
-1	1	1	1	-1	-1	0.933	0.907	0.913	0.913	0.880	0.923	0.890	0.897
1	-1	1	1	1	1	1.023	1.030	1.030	1.130	0.973	1.000	1.013	0.997
-1	1	-1	1	1	1	1.290	1.220	1.270	1.157	1.280	1.223	1.170	1.207
1	1	1	-1	1	-1	1.100	1.033	1.017	1.007	1.017	1.053	1.000	1.090
1	-1	-1	1	-1	-1	0.890	0.850	0.850	0.850	0.860	0.930	0.870	0.810
-1	-1	1	-1	-1	1	11.730	7.560	10.400	12.070	24.870	12.870	18.750	
1	1	-1	-1	-1	1	16.870	8.730	11.070	14.830	16.400	19.630	13.470	14.970
-1	1	1	-1	1	-1	1.273	1.230	1.263	1.247	1.263	1.250	1.217	1.167
-1	-1	1	1	1	1	1.003	0.970	0.973	0.967	0.970	0.923	0.907	0.867
-1	-1	-1	1	-1	-1	0.907	0.907	0.883	0.823	0.837	0.873	0.890	0.867
-1	1	-1	-1	-1	1	4.000	6.700	8.000	2.700	8.200	4.000	8.200	4.200
1	1	-1	1	1	1	1.140	1.117	1.090	1.073	1.117	1.097	1.157	1.057
1	-1	-1	-1	1	-1	0.937	0.960	0.953	0.873	1.017	0.927	0.937	0.910
1	1	1	1	-1	-1	0.840	0.853	0.863	0.863	0.823	0.783	0.790	0.817
1	-1	1	-1	-1	1	26.070	18.970	19.600	18.700	25.300	16.500	18.630	26.530
-1	-1	1	-1	-1	-1	2.003	2.030	2.490	4.457	3.440	3.007	2.473	4.517
-1	-1	-1	-1	1	1	1.073	1.003	1.073	1.087	1.090	1.090	1.073	1.047
1	-1	-1	1	-1	1	0.780	0.780	0.863	0.897	0.823	0.763	0.830	0.807
-1	1	1	1	-1	1	0.907	0.893	0.840	0.970	0.863	0.823	0.823	0.850
1	-1	1	1	1	-1	0.977	1.087	1.077	1.140	1.060	1.083	1.087	1.140
1	1	-1	-1	-1	-1	1.430	1.553	1.050	1.197	1.450	1.303	1.173	1.707
1	1	1	-1	1	1	1.067	1.147	1.107	1.123	1.097	1.040	1.050	0.963
-1	1	-1	1	1	-1	1.090	1.000	1.030	1.060	1.050	1.030	1.040	1.050

Table 2. The raw data matrix. r1 - r8 are the repetitions.

Model: CONSTANT+A+B+C+D+E+F+AB+AC+AD+AE+AF+BC+BD+BE+BF+CD+CE+CF+DE+DF+EF+ABC+ABD +ABE+ABF+ACD+ACE+ACF+ADE+ADF+AEF+BCD+BCE+BCF+BDE+BDF+BEF+CDE+CDF+CEF+DEF						
Dependent variable	original response λ=1	St.d/mean:	0.106	transformed response λ=-1	St.d/mean :	0.050
Effect	t	t	P(2 Tail)	t	t	P(2 Tail)
CONSTANT	2.025	9.432	0.000	0.415	19.823	0.000
A	0.321	1.497	0.149	-0.015	-0.732	0.472
B	0.210	0.978	0.339	0.011	0.526	0.604
C	**0.768**	**3.577**	**0.002**	0.028	1.314	0.202
D	**-2.055**	**-9.570**	**0.000**	**-0.513**	**-24.400**	**0.000**
E	**-1.976**	**-9.205**	**0.000**	**-0.325**	**-15.500**	**0.000**
F	**1.677**	**7.809**	**0.000**	**0.146**	**6.970**	**0.000**
AB	-0.167	-0.779	0.445	-0.026	-1.257	0.222
AC	-0.187	-0.872	0.393	0.008	0.358	0.724
AD	-0.343	-1.596	0.125	-0.040	-1.894	0.071
AE	-0.318	-1.481	0.153	0.022	1.067	0.297
AF	0.364	1.697	0.104	0.028	1.329	0.198
BC	0.235	1.095	0.286	-0.024	-1.166	0.256
BD	-0.203	-0.947	0.354	0.004	0.186	0.855
BE	-0.200	-0.930	0.362	0.011	0.512	0.614
BF	0.210	0.979	0.338	0.032	1.525	0.142
CD	**-0.789**	**-3.676**	**0.001**	**-0.076**	**-3.635**	**0.001**
CE	**-0.796**	**-3.709**	**0.001**	**-0.087**	**-4.162**	**0.000**
CF	**0.543**	**2.530**	**0.019**	**-0.086**	**-4.123**	**0.000**
DE	**2.025**	**9.429**	**0.000**	**0.451**	**21.529**	**0.000**
DF	**-1.669**	**-7.775**	**0.000**	**-0.133**	**-6.371**	**0.000**
EF	**-1.639**	**-7.631**	**0.000**	**-0.068**	**-3.262**	**0.004**
ABC	-0.176	-0.819	0.421	-0.002	-0.103	0.919
ABD	0.163	0.760	0.455	0.025	1.186	0.248
ABE	0.141	0.659	0.517	-0.021	-1.012	0.322
ABF	-0.096	-0.448	0.659	0.030	1.409	0.173
ACD	0.204	0.951	0.352	0.028	1.339	0.194
ACE	0.200	0.933	0.361	0.029	1.395	0.177
ACF	-0.161	-0.748	0.462	-0.011	-0.545	0.591
ADE	0.332	1.546	0.136	0.024	1.154	0.261
ADF	-0.371	-1.728	0.098	**-0.044**	**-2.112**	**0.046**
AEF	-0.350	-1.629	0.118	-0.003	-0.123	0.903
BCD	-0.274	-1.275	0.216	**-0.069**	**-3.317**	**0.003**
BCE	-0.240	-1.119	0.275	0.009	0.424	0.676
BCF	0.212	0.987	0.335	0.001	0.063	0.950
BDE	0.195	0.907	0.374	-0.030	-1.417	0.171
BDF	-0.184	-0.858	0.400	0.038	1.793	0.087
BEF	-0.210	-0.977	0.339	-0.024	-1.169	0.255
CDE	**0.780**	**3.631**	**0.001**	**0.051**	**2.435**	**0.023**
CDF	**-0.563**	**-2.622**	**0.016**	0.041	1.938	0.066
CEF	**-0.578**	**-2.692**	**0.013**	0.021	1.000	0.328
DEF	**1.662**	**7.740**	**0.000**	**0.121**	**5.801**	**0.000**
Std Error	**0.215**			**0.021**		
Analysis of Variance	original response λ=1	Squared multiple R:	0.965	transformed response λ=-1	Squared multiple R:	0.986
Source	Sum-of-Squares	F-ratio	P	Sum-of-Squares	F-ratio	P
Regression df=41	1809.594	14.958	0.000	43.294	37.617	0.000
Residual	64.915			0.618		

Table 3. ANOVA output original data and inverse data. Bold typing indicate significance (p<0.05)

D	E	F		λ=-3	λ=-1	λ=0	λ=1
-1	-1	-1			2.088	2.254	2.435
-1	-1	1		2.323	12.081	13.961	15.728
-1	1	-1			1.027	1.030	1.033
-1	1	1		1.059	1.114	1.119	1.125
1	-1	-1			0.928	0.933	0.938
1	-1	1		0.906	0.900	0.903	0.906
1	1	-1			0.984	0.986	0.988
1	1	1		1.006	1.041	1.045	1.049
95% interval at							
-1	-1	1	Upper	Not Available	44.168	15.333	17.189
			Lower	1.666	6.997	12.711	14.267
1	-1	1	Upper	0.939	0.951	0.992	2.367
			Lower	0.877	0.853	0.822	-0.555

Table 4. Estimated expectations and 95 % confidence intervals. For λ=-3 only D, E and DE are significant and are included in the model. The upper confidence limit for λ=-3 when D=-1 and E=-1 is not available because the value is out of range on the original scale.

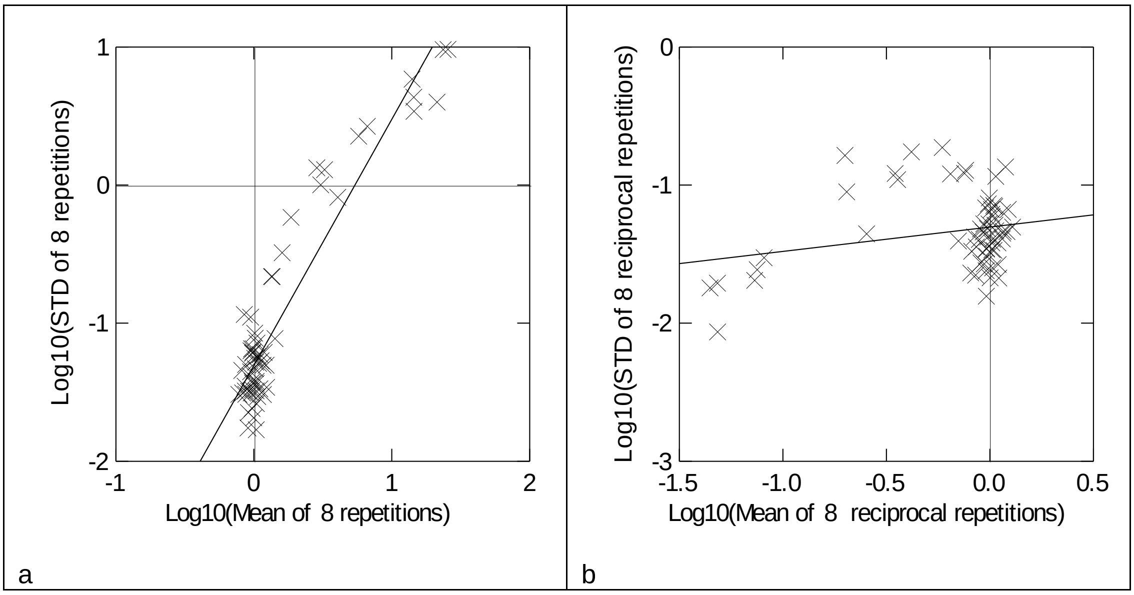


Figure 1. Log(Mean) against log(STD) of repetitions at the 64 design points. Frame a: the original data, frame b: the reciprocal transformed data.

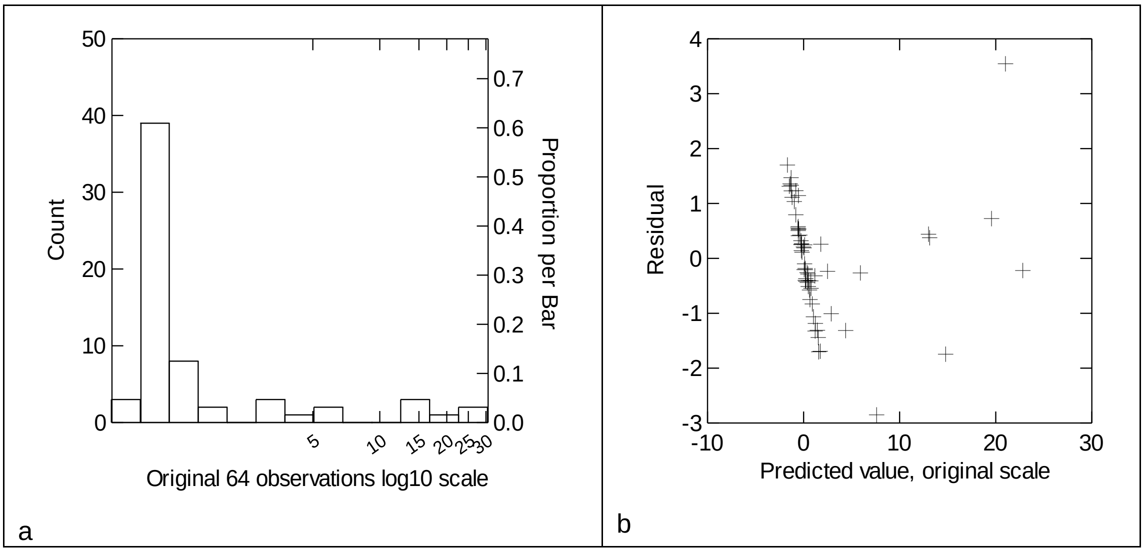
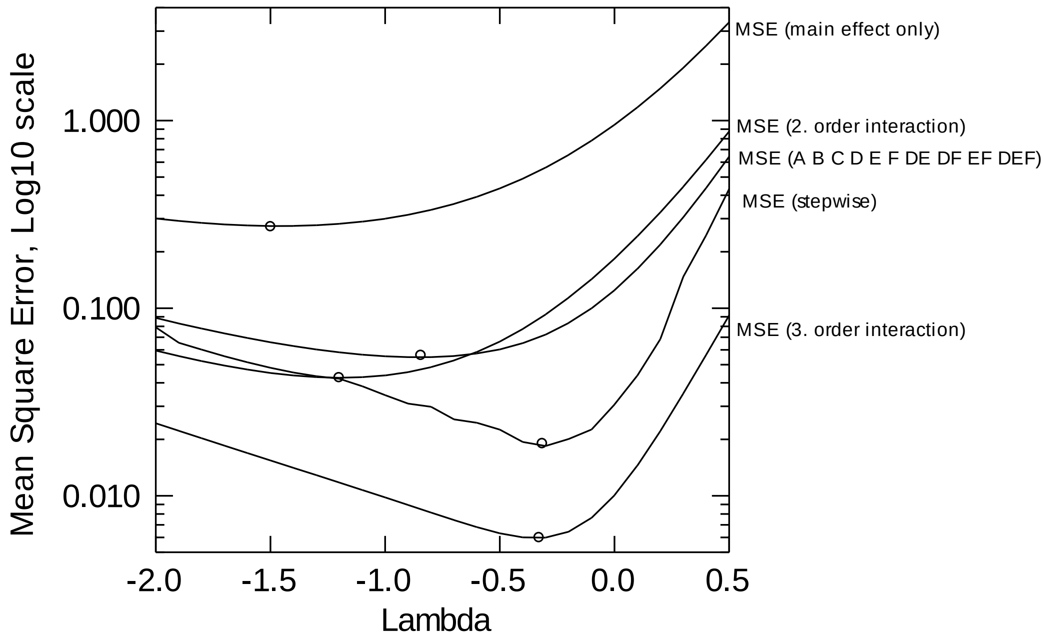


Figure 2. Frame a:The original 64 observations. Frame b: Residuals after fitting the model



The mean square errors (MSE) are shown on the vertical axis (using a Log_{10} scale in order to include several curves) from the five different regression models, and with a range of values of λ on the horizontal axis.
◦ indicate sites of minimum MSE values

Figure 3. Mean squares errors from five models.

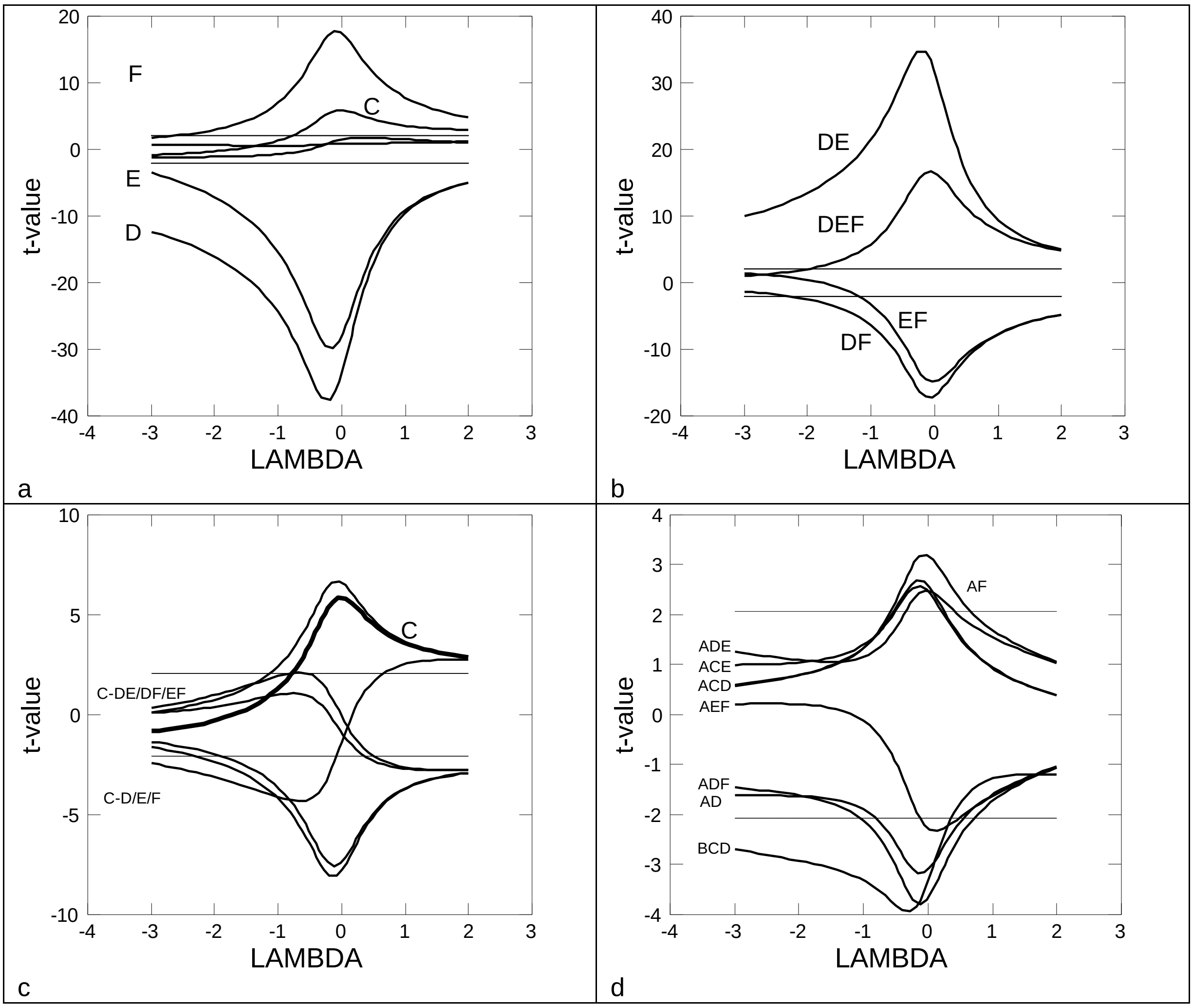


Figure 4. t-values of effects against lambda. Regression model no 3. Frame a: the main effects. Frame b: the D, E, and F interactions. Frame c: the interactions between C and D, E and F. Frame d: some other interactions. . The dashed lines indicate 5% significance.

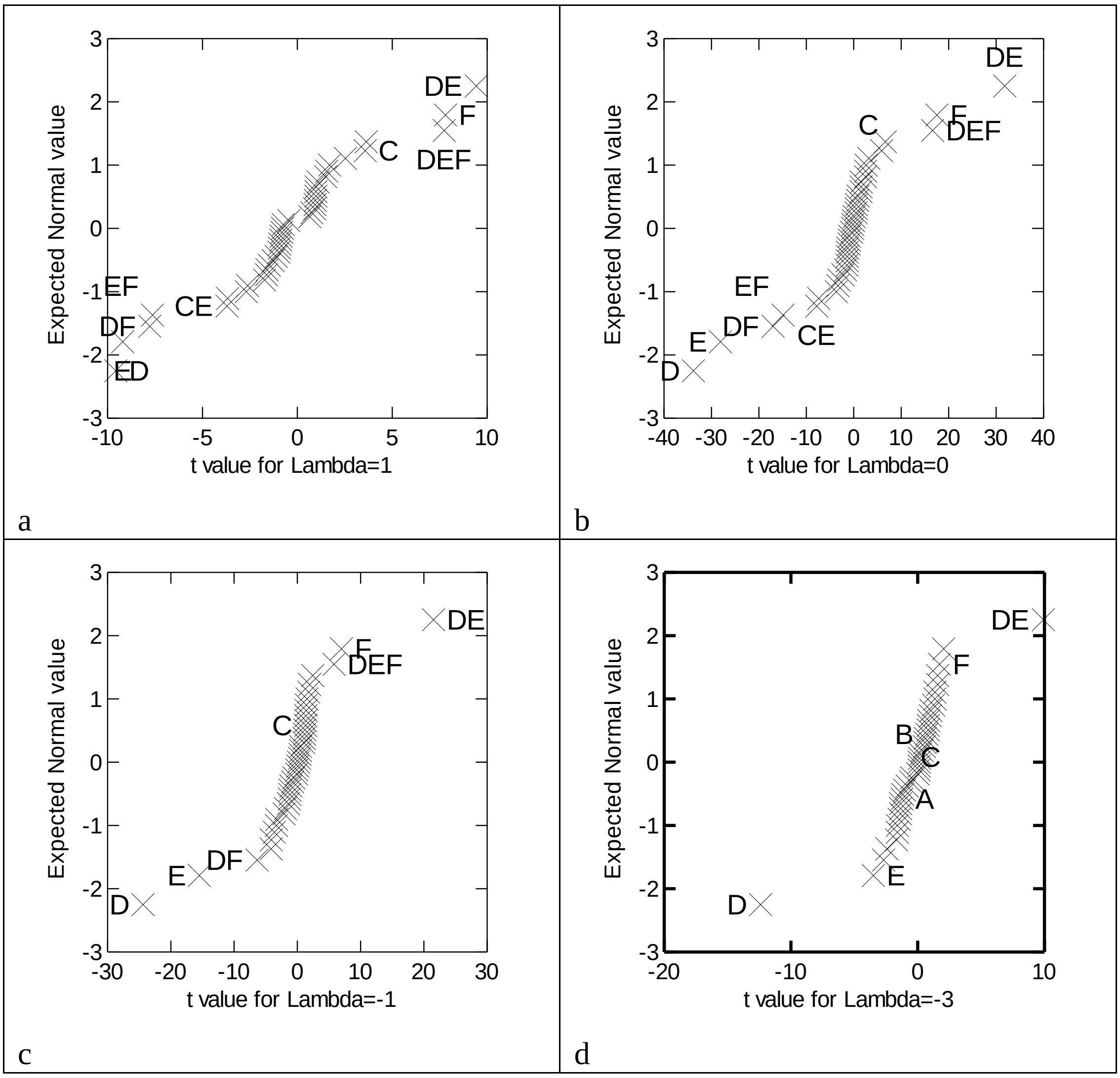


Figure 5. Normal probability plots of estimated effects. Regression model no 3. Frame a-d are for selected λ values.

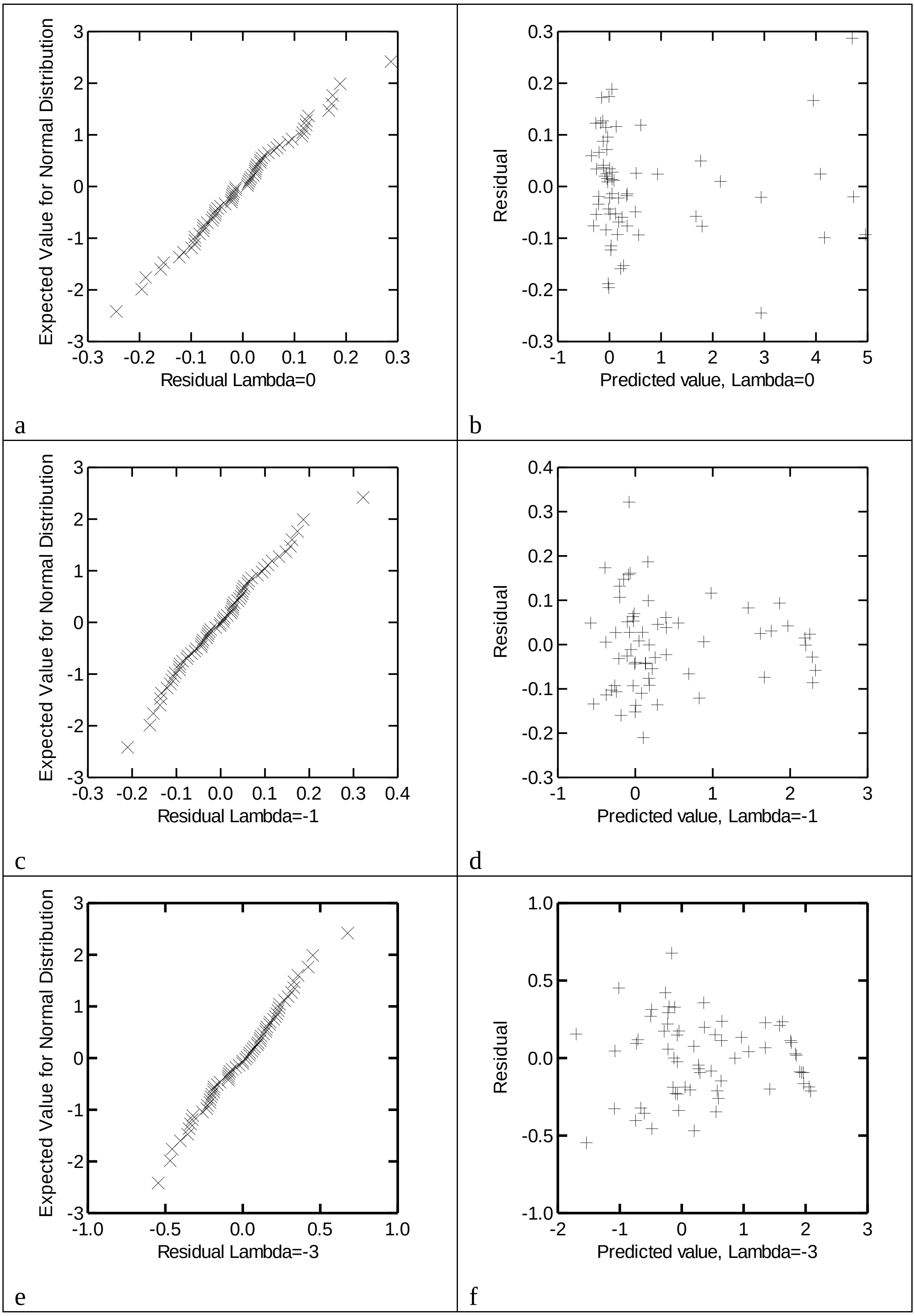


Figure 6. Residuals after fitting a model with up to 3-factor interactions. Regression model no 3. Frame a, c and d: Residuals, λ = 0, -1 and -3. Frame b, d and f: Predicted values, λ = 0, -1 and -3.

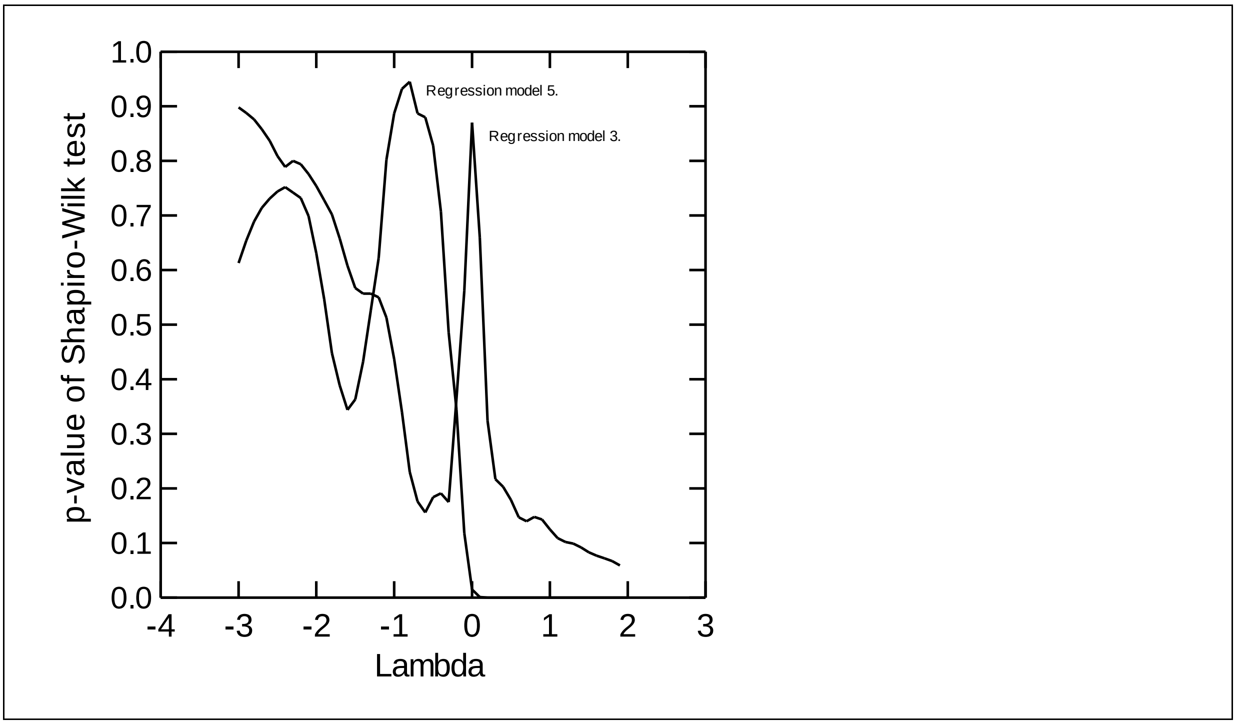


Figure 7. Shapiro-Wilk test for normality. The p-value for different λ and the regression models no 3 and 5.